



Changing neighborhoods, shifting connections: mapping relational geographies of gentrification using social media data

Ate Poorthuis ^a, Taylor Shelton ^b and Matthew Zook ^c

^aDepartment of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium; ^bDepartment of Geosciences, Georgia State University, Atlanta, United States; ^cDepartment of Geography, University of Kentucky, Lexington, United States

ABSTRACT

New sources of ‘big data’ are regularly described as revolutionizing the study of urban life. Of particular interest is analyzing gentrification, which has proven a challenging endeavor with conventional methods. Big data may offer a new approach to the persistent problem of defining and measuring gentrification, while also allowing us to rethink broader questions about theory and methodology in urban geography. Using geotagged Twitter data, we demonstrate how the changing geographies of users’ tweets are proxies for the evolving social and spatial contours of urban neighborhoods. We use the case of Lexington, Kentucky to analyze the mobilities and relational connections of neighborhood residents and visitors as gentrification intensified over time. We argue that these kinds of big data allow for an analytical approach that focuses on the dynamic, relational connections between people and places, and provides a useful, additional avenue in understanding a process as complex and multifaceted as gentrification.

ARTICLE HISTORY

Received 30 July 2020
Accepted 4 February 2021

KEYWORDS

Gentrification; big data; spatial networks; social media; Twitter; Lexington, KY; relational geography

I. Cities in the age of big data

The emergence and rapid growth of new sources of so-called “big data” are widely thought to have revolutionized the study of social life. This is especially true for our understanding of cities, where nearly every aspect of social interaction is captured and stored, whether through CCTV, sensors embedded in streets, sidewalks and buildings, smartcard systems, social media or mobile phone tracking. These new data sources are particularly useful insofar as they allow for more a spatially and temporally granular approach to analyzing urban social and spatial processes than is possible with more conventional datasets. This promise, however, has led to the over-valorizing of big data as the means to address all urban problems. Regardless of the particular context or subject of a dataset, the prevailing discourse around big data is predicated on the “widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy” (boyd and Crawford, 2012, p. 663).

Moreover, not all aspects of the city or urban life are equally quantified or datafied. Reasons for a relative dearth of data can range from a national statistical agency not collecting or not publicly sharing data due to privacy concerns, or simply because it would be too costly in both time and resources to collect. Other key aspects of urban life are hard to reduce to numbers at all, such as sense of place or community. Many processes, however, fall somewhere in between. They may lend themselves to datafication, but appropriate data is not collected or made easily accessible. Of course, even where data is available, it still provides only a partial picture of the process at hand. And given the dictum that “what gets measured gets managed”, the areas for which data is less available or less robust often become significant blind spots in attempts to intervene in urban processes. Gentrification is a particularly good example of a phenomenon that has long defied easy measurement, complicating efforts to understand and respond to it.

Despite its position as a cornerstone of urban research for the last half century, gentrification’s place in broader public consciousness is more recent, growing rapidly over the past decade. Debates over whether gentrification is a bad thing – indeed, whether it even exists in the first place – have led some to question its utility, calling it a “vague, imprecise and politically loaded term ... [in] need [of] better, more objective ways to measure it” (Florida, 2014). While our paper doesn’t question the importance of gentrification as a concept or the importance of the rich body of critical scholarship about it, we do agree that there has been a persistent problem of defining and measuring it. However, the advent of big data offers an opportunity to rethink broader questions about theory and methodology in geography and GIScience (Poorthuis & Zook, 2020), while also addressing the specific tendency of gentrification research to eschew methodological discussions (Slater et al., 2004).

As such, this paper explores the potential for big data – in particular, geotagged social media data from Twitter – to analyze gentrification. We focus especially on the use of geotagged Twitter data as an indicator of social activity and relations over time and space, thus serving as a proxy for the evolving social and spatial contours of urban neighborhoods. Building on earlier work demonstrating the utility of this data to study the aggregate mobility patterns of urban residents (Shelton et al., 2015), we seek to understand gentrification as a reconfiguration of the mobilities and relational connections between urban neighborhoods.

We use a case study of Lexington, Kentucky and the city’s gentrifying Northside to analyze how the everyday mobilities of Lexington residents in these neighborhoods have evolved as gentrification intensified over time. We identify and visualize these changes in neighborhood dynamics by representing the fundamentally relational and mobile lives of urban residents through spatial network analysis, in contrast to an exclusive focus on residential geographies or the internal characteristics of neighborhoods.

Given the relative novelty of this data and method, it is crucial to note from the outset that we do not argue that this approach represents the only, or even necessarily the best, way to understand gentrification. As detailed below, this approach still misses multiple elements of the gentrification process, such as the particular mechanisms of reinvestment and displacement and the cultural politics of displacing racial and ethnic minorities from neighborhoods they have long called home. Instead, we provide an additional avenue for understanding the complex and multifaceted process of gentrification, focused on how gentrification is produced through dynamic, relational connections between people and

places within the city. To frame our approach, we review (1) how gentrification has been identified and analyzed using more conventional datasets, (2) the development of relational understandings of place, geographies and gentrification, and (3) other nascent attempts to measure gentrification and associated processes of urban inequality using big data. With this grounding we then turn to an empirical analysis of geotagged Twitter data in Lexington, Kentucky to demonstrate the utility of our approach.

II. Understanding the geographies of gentrification and urban change

Using quantitative data to understand cities is by no means new. Nor is the ostensibly more objective and scientific air around the use of such methods. For the past century or more, urban planners, policymakers and scholars have used quantitative data under the goal or guise of constructing a scientific approach to cities, even in the absence of “big” data (cf. Barnes, 2013; Barnes & Wilson, 2014; Fairfield, 1994; Ford, 1913; LeGates et al., 2009; Light, 2003; Shelton, 2017; Zook, 2017). This has been no less true for studies focused specifically on the process of gentrification, defined as the transition of neighborhood character that accompanies the out-migration of previous, and generally poorer, residents who are displaced by an influx of newer, and generally richer, residents. While gentrification has been a topic of interest to geographers, sociologists and planners for more than 50 years (Glass, 1964), it has remained a much-debated and contested concept, owing in large part to the fundamental racial and class inequalities to which it calls attention.

A key part of this debate, and indeed a source of much disagreement, is the lack of an agreed-upon method for measuring, identifying or analyzing gentrification. This section reviews how geographers and other urban social scientists have sought to measure gentrification quantitatively, using both conventional data as well as more recent iterations of big data, and how these approaches to defining and measuring gentrification intersect with theoretical trends around relational conceptualizations of space.

Defining and measuring gentrification

Gentrification has long been considered a “chaotic concept” (Rose, 1984), a fact that has only been exacerbated by more recent public attention to the process and its implications in cities around the globe (cf. Buntin, 2015; Cortright, 2015; Economist, 2015; Florida, 2014). But, as Hammel and Wyly (1996) have argued, “[u]ncertainty over the extent of gentrification stems not only from the complexity of the process, but also from the difficulty of observing and measuring the phenomenon” (Hammel & Wyly, 1996, p. 248). That is, the problem isn’t just a lack of agreement on what gentrification is, but that our agreed-upon definitions often don’t match up with the data used to measure and track gentrification as it unfolds. Nonetheless, geographers and urbanists have spent considerable energy on attempts to quantitatively measure gentrification processes using a variety of data sources. As such, there are three key ways that gentrification is defined, and thus measured.

First is the notion of gentrification as a class, or more generally demographic, transition within a neighborhood. In other words, gentrification is defined as a change in the

makeup of the people who live in a particular area. Given the relative accessibility of data on the demographic makeup of households and neighborhoods (proxied through census tracts or other administrative units) with regards to income, race, educational attainment and so on, many quantitative empirical studies of gentrification rely on this understanding in order to meaningfully operationalize it. While income is perhaps the most significant variable for understanding changes in the class composition of a given neighborhood, income alone fails to capture the variety of ways that these changes are expressed. For example, scholars have often relied on the percentage of residents with a college education (Freeman, 2005; Hammel & Wyly, 1996; Schuler et al., 1992) or in professional careers (Atkinson, 2000) as alternative ways of measuring these changes, acknowledging that many gentrifiers are younger people at the beginning of their earning careers who do not actually make all that much money. While the use of race and ethnicity variables has been somewhat less common, the current context of US cities means that gentrification often unfolds in inner city areas with larger black and Latino populations, making race a key part of the practice and experience of gentrification (Kirkland, 2008), although this isn't always the case in contexts beyond the US.

The second major understanding of gentrification focuses on changes to, and investments in, the material, built environment of a neighborhood. In diverting focus from the people moving into a neighborhood, such an understanding of gentrification resonates with Smith's (1979) classic characterization of gentrification as "a back to the city movement by capital, not people". While this vein of work sometimes relies on qualitative data taken from field surveys to assess visual indicators of building upgrading (cf. Wyly & Hammel, 1998, 1999; Hammel & Wyly, 1996 for a combination of such a field survey with Census data, or Hwang & Sampson, 2014, who use Google Street View imagery as a way of performing such an analysis remotely), some quantitative indicators also serve as reasonable proxies for such upgrading. These include changes in rents, property valuation and the proportion of owner-occupied houses (Heidkamp and Lucas 2006), as well as increased mortgage lending (Kreager et al., 2011; Wyly & Hammel, 1999).

Finally, there is the crucial notion that for demographic change and material upgrading to count as gentrification, there must also be *displacement*. Slater et al. (2004) call attention to the fact that "[d]isplacement is vital to an understanding of gentrification, in terms both of retaining definitional coherence and of retaining a critical perspective on the process" (1144), but this concept tends to be absent from quantitative approaches to measuring gentrification. This is because displacement, despite its role as one of the central aspects of gentrification, is also the hardest aspect to measure (Easton et al., 2020). While secondary data from government sources can show demographic changes within a given neighborhood that are consistent with displacement (e.g. shifts from poorer and nonwhite residents to higher income and whiter households), it still tells a very limited story, as the reasons behind these particular changes are not captured in such data. As Carlson (2020) shows, such population-based measures actually diverge quite significantly from other, more granular measurements of displacement that account for individual-level residential mobility and the reasons for such mobility.

A more recent arrival within gentrification research is the use of non-official – but not quite "big" – data sources to examine other aspects of the processes. These sources can be quite diverse, ranging from the local knowledge of activists and nonprofit groups about neighborhood conditions contained to newspaper articles referencing gentrification.

Given the differences in data sources and implicit definitions of gentrification, it is unsurprising that the findings of this work range from more conventional gentrification research, as in Barton's (2016) comparison of neighborhoods identified as gentrifying in *New York Times* articles to earlier quantitative work by Bostic and Martin (2003) and Freeman (2005) that use Census data to track neighborhood-scale demographic changes. Other studies have turned to more cultural markers of gentrification, focusing on spaces of consumption such as the number of cupcake and coffee shops to track shifts in neighborhood character (Papachristos et al., 2011; Smith, 2014; Twilley, 2009, 2011). With the increasing digitalization of everyday social life, traces of these changes are more and more visible in online spaces, such as in Yelp restaurant reviews, which Zukin et al. (2017) use to show the impact of racialized spatial perceptions on gentrification processes. While these approaches do not focus on the more conventional political-economic and structural character of the gentrification process, they are important insofar as they capture "an on-the-ground and visible manifestation of a particular form of gentrification – the increased presence of an amenity often associated with gentrifiers' lifestyles" (Papachristos et al., 2011, p. 216). But using these kinds of ever-changing cultural markers as indicators of gentrification poses a temporal problem that limits the replicability of research, as cupcake shops and coffee shops may lose their cachet in gentrifiers' changing cultural tastes and consumption habits.

Therefore, the persistent challenge in quantitative gentrification research is finding an alignment between one's definition of gentrification as a concept and the available data to measure the process empirically. While most studies of gentrification have operationalized their definition in largely territorial and spatially contiguous terms, i.e. the neighborhood, we argue that it is fundamental to also include a relational understanding of space and socio-spatial processes. This relational grounding of gentrification also opens up opportunities for utilizing big data to meaningfully contribute to the study of changes in urban life.

Understanding urban change relationally

For the last four decades, critical human geography has conceptualized space and place not as fixed, unchanging containers of social activity, nor an external force on social life, but rather as dynamically produced social products that in turn help to shape social processes (cf. Soja 1980; Lefebvre 1991). Inspired in large part by the work of Doreen Massey (1991), this relational understanding of space emphasizes that space is not preexisting but is only produced through relationships between people and places.

In contrast, Tobler's "first law of geography" that "all things are related, but near things are more related than far things" is often used in quantitative geography as an illustration of a more conventional, absolute understanding of space. However, such conceptualizations of space, including those by quantitative geographers, have long included more nuance than Tobler's first law suggests (Poorthuis & Zook, 2020). For example, Tobler himself notes in his second "law", that "the phenomenon external to an area of interest affects what goes on inside". This parallels a relational understanding in which no space or place forms an entirely coherent whole detached from broader flows and processes that constitute the social world. All spaces and social processes – including

gentrification – are produced through relations that are stretched across space and time to lesser and greater extents.

Building on this work, we argue for an additional, alternative definition of gentrification. That is, we see gentrification as a fundamentally relational process that reconfigures how particular people are connected to particular places (see also Lawton, 2020). Such a definition of gentrification provides room for an understanding that reconfigurations often take the form of the wealthy displacing the poor from residential spaces, while also acknowledging that other kinds of spatial relationships are also reshaped through gentrification. Just as Smith's notion of a back to the city movement by capital can be tracked via the financial flows made visible by mortgage lending data or property records, gentrification's reconfiguration of the way people interact with urban neighborhoods and places might be traced via big data sources such as social media.

Of course, understandings of relationality have always been implicit in gentrification research. The idea that places are not static is, in many ways, a fundamental precept of gentrification as process, whether one conceives of that process as being primarily due to the changing demands of capital or the changing consumption preferences of the middle and upper classes. But a paucity of data that captures this relationality has caused such an understanding to recede into the background of such work. In the US case, the Census builds from a constitutional mandate of decennial population counts to define *contiguous* congressional districts and this conceptualization of space is perpetuated throughout social research designs and data structures. Even data specifically focused on mobility, *e. g.* how people move around a city, is often limited to commuting relations or travel surveys, which are collected infrequently and for a relatively small samples, insufficient for neighborhood-level studies.

At the same time, there is a tendency in critical scholarship on gentrification to view gentrifying neighborhoods as hermetically-sealed containers that should stay that way, which has limited the take-up of relational conceptualizations. But understanding gentrification as a fundamentally relational and spatially extensive process can actually help overturn the stigmas placed onto newly-gentrifying areas. That is rather than blaming residents for the current and historical problems of the neighborhood, these issues are understood as produced by broader social and spatial processes that go beyond the boundaries of the neighborhood (*cf.* Massey 1979 on the ways that “inner city” or “regional” problems are explained through characteristics internal to such areas).

How big data can change gentrification research

From this foundation of differences and debates on measuring and conceptualizing gentrification, big data provides both clarity and confusion to the existing scholarship. As alluded to earlier, big data might provide hitherto difficult-to-collect data on people's connections to urban places and neighborhoods. At the same time, the operationalization of a relational understanding of gentrification has the potential to exacerbate the messiness with which gentrification is analyzed and mobilized discursively. Thus, for us, the challenge of using big data in gentrification research is to clearly understand how this data adds something to our existing understandings of gentrification. That is, the fact that data has long been used to measure and map gentrification points to the reality that data

can't "solve" the problem of defining gentrification and its political, economic and moral implications.

Attempts at mobilizing big data, and especially geotagged social media data, to understand gentrification remain in their infancy, but represent a dynamic and growing field of geography. So far, however, this work has tended to focus more on novelty than on substantively advancing knowledge of the multi-faceted gentrification process. For example, using geotagged tweets, Foursquare check-ins or OpenStreetMap data, as Beekmans (2011), Schaefer (2014), and Venerandi et al. (2015) do, is certainly innovative, but these approaches typically fall back on comparing the densities of data of different types in different areas. Places with more tweets or more Foursquare check-ins associated with a given gentrification-related topic, such as tweets about "yuppies" or "eviction", are given a higher value and interpreted as indicative of the presence of gentrification. But relying on such an approach replaces one partial dataset with another. Indeed, as Schaefer's (2014) research shows, relying on collecting tweets for gentrification-related keywords yields only a few thousand data points from which to work, even in a city as large as Los Angeles, suggesting that such an approach would be of limited utility.

A more productive route, we argue, is leveraging the relationality inherent in this data as a means to measure the relationality inherent in the gentrification process. Arguably the greatest contributions that big data make to geographic research more generally is that it can help overcome persistent barriers in terms of the spatial, temporal and relational resolution of data. Census data is collected on a relatively infrequent basis and aggregated to areal units that are often fairly coarse, especially for understanding something like gentrification (Schuler et al., 1992). As gentrification processes typically unfold faster than the decennial census, such data can be ill-suited to identify and analyze changes in time. Moreover, gentrification regularly occurs at spatial scales that differ from the standard Census tracts creating conflicts between the vernacular and administrative definitions of the neighborhood. Indeed, gentrification actively reshapes the way neighborhoods are thought about and named by residents (Madden, 2018). Big data provides an avenue for constructing different areal units as needed and using geotagged social media data as a proxy for the mobility patterns of populations (Blanford et al., 2015; Luo et al., 2016; Phillips et al., 2019; Wu et al., 2016; Zhang et al., 2016). These mobilities allow us to develop alternative definitions of neighborhood boundaries and their relational connections through the movements of people that are connected to such places (cf. Poorthuis, 2018; Shelton & Poorthuis, 2019; Shelton et al., 2015). Such an approach parallels relational understandings of space in that it acknowledges that people are fundamentally mobile and connected beyond immediate spatial proximity and that these patterns say as much about us as our places of residence (Kwan, 2012, 2013).

In the analysis that follows, we combine these approaches, opting for a more relational understanding of urban space as seen through the everyday mobilities of Twitter users, while also relying on more traditional neighborhood definitions and statistical geographies in order to leverage additional social data that would otherwise be unavailable at this more flexible and relational geography.

III. Relational geographies of gentrification in Lexington, KY

As a proof of concept of this relational conceptualization of gentrification operationalized through big data, we analyze the case of Lexington, Kentucky (see [Figure 1](#)), a medium-sized

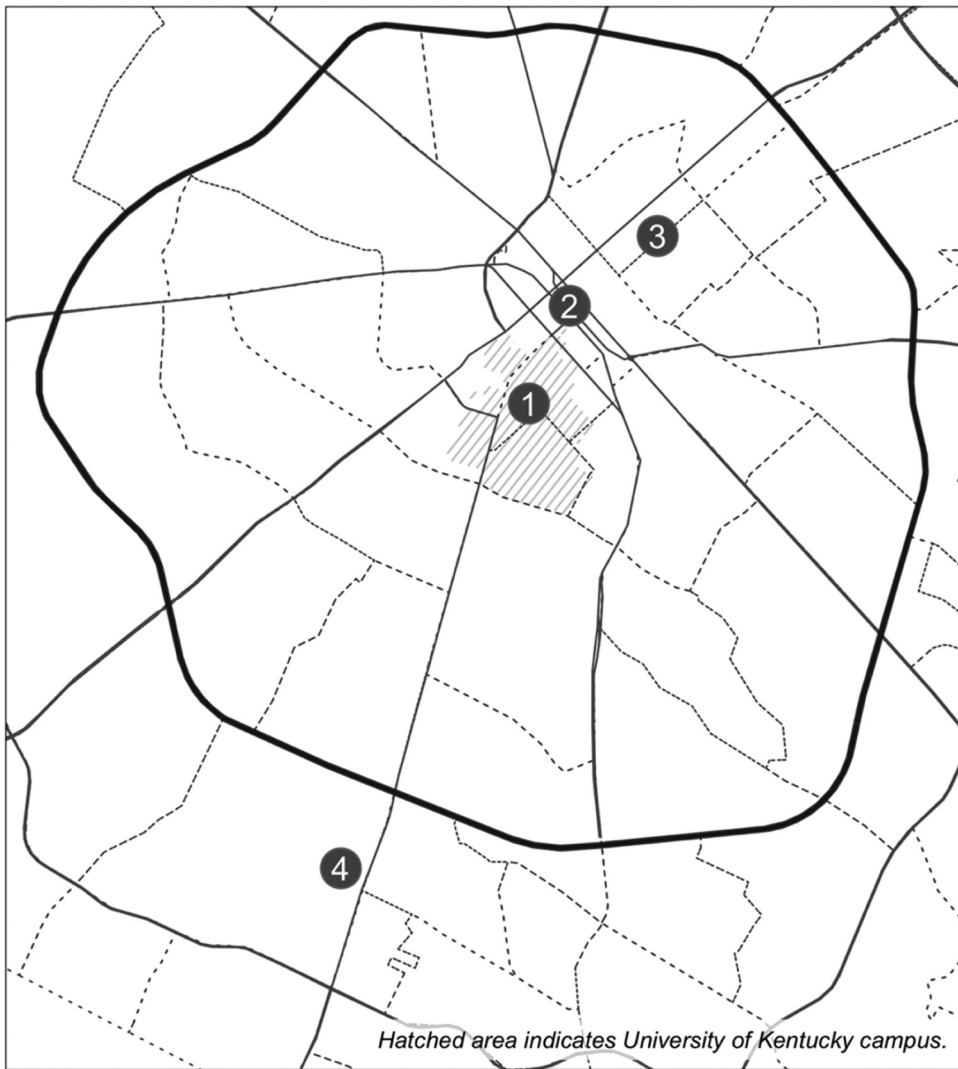


Figure 1. Map of study area, Lexington, KY. Labeled places are (1) University of Kentucky; (2) Downtown; (3) Northside neighborhood; and (4) Fayette Mall.

city in the southern United States with a core population of roughly 300,000 and a surrounding hinterland of an additional 200,000 people. While our case selection is driven in part by a desire to understand gentrification beyond the “usual suspects” of New York, Los Angeles and other larger cities, another consideration is our familiarity with the city, allowing us to “ground truth” our quantitative analysis with an understanding of the specifics of the gentrification processes as they’ve unfolded over the last decade. This is an important nuance in the use of big data for studying gentrification: we see such data as being best employed *in addition to* existing methodologies – including qualitative fieldwork-based methods – and not as a be-all-end-all replacement for more established approaches.

Our analysis is based on a dataset of all geotagged tweets produced within the greater Lexington, KY¹ area between June 2012 and December 2017 extracted from the DOLLY

system at the University of Kentucky (Poorthuis & Zook, 2017). During this period, 95,285 different users sent a total of 4.6 million geotagged tweets from the Lexington area. In order to ensure that we had sufficient data to measure changes in individuals' movements and activities we eliminated any users with fewer than 10 tweets, leaving a total of 4.4 million tweets by roughly 25,000 users.

Previous work with geotagged tweets (and other social media) has often focused on the location of each tweet – essentially treating each data point in isolation and subsequently aggregating the total number of data points per spatial unit. While appropriate for many research questions, treating each data point separately leaves considerable information unused. Other work (cf. Poorthuis, 2018; Shelton & Poorthuis, 2019), has drawn connections between data points generated by the same user to associate different locations with each other. In this analysis, we take a slightly different approach in drawing connections between a user's data points. Namely, we determine the most significant or “home” location for each user and draw links between this “home” location and other locations within the city visited by that user. “Home” is written between quotation marks² because our intent is not to represent the actual, “true” home location of the user but rather the most significant base location for that user (specifically defined as census tract)³ based on the following criteria:

- (1) The census tract has to be tweeted from at least five times by the user;
- (2) The user has to have tweeted from that census tract on at least five separate days;
- (3) The earliest tweet and latest tweet from that census tract are at least ten days apart.

This approach, including a comparison with different algorithms for “home” location inference, is discussed in more detail by Chen and Poorthuis (2021). These criteria eliminate individuals with only a transient presence both digitally and materially, whether because the account wasn't used consistently over time, or the user was tweeting while visiting Lexington (and therefore is not representative of the evolution of intra-urban mobility dynamics in Lexington over time). In the event that more than one location matches these conditions for a given user, we use the location with the most tweets as the user's home location (in case of a tie, we randomly pick one location).

Using this heuristic, we are able to determine a home location for 17,476 users who collectively sent 3.8 million tweets during the study period. The subsequent analysis is based on two key features that we refer to as *visits* and *profiles*. A *visit* is defined when a user (a visitor) sends a tweet from a Census tract other than their home location and by doing so establishes a connection between the two locations. *Profiles* are based on the characteristics of the home census tract for each user. We know relatively little about each user solely from their tweets. Therefore, we rely on what is known about their home location from the Census, such as racial composition, median household income and educational attainment, to create potential demographic profiles of users based in such neighborhoods, with the important caveat that this does introduce an ecological fallacy. To address the presence of power users (who tweet much more than most users), we only count one *visit* per year (regardless of how many additional tweets were made in that location by the user).

It should be noted that the Twitter users in this dataset do not represent a representative sample of the population of Lexington, just as they are not representative

of the population at large (Hargittai, 2015). As such, the subsequent analysis is not an attempt to provide *the* single, authoritative view on gentrification in the city, but rather to add an additional perspective to our understanding of gentrification and how it might be measured. Specifically, this additional view highlights insights we might gain from a relational approach that accounts for residents' everyday mobilities within the city and across different neighborhoods, rather than being limited to the more conventional, residentially-oriented datasets used for this type of analysis.

Stability and change in Lexington's intra-urban mobilities

We first test whether the home location of users is stable over the entire 5.5-year study period. To do so, we apply the home location heuristic for two separate periods: (1) June 2012 to December 2013 and (2) January 2014 to December 2017⁴. As can be seen in Figure 2b, the amount of "migration" between the two periods is relatively minor with only 11% of users (1,958 out of 17,476) assigned a different home location in Period 2. In other words, the migration of users' home locations alone doesn't represent a meaningful indicator for understanding the gentrification process. Given the relative stability of the home location, we elected to use the home location based on the entire study period (see Figure 2a) throughout the analysis to maximize the available data per user.

Examining the connection between users' home location and visits within the city (Figure 3) highlights a number of patterns. The three most popular tracts for visits contain the campus of the University of Kentucky, as well as neighborhoods popular for student housing (the dark brown tract between the two orange tracts in Figure 3).

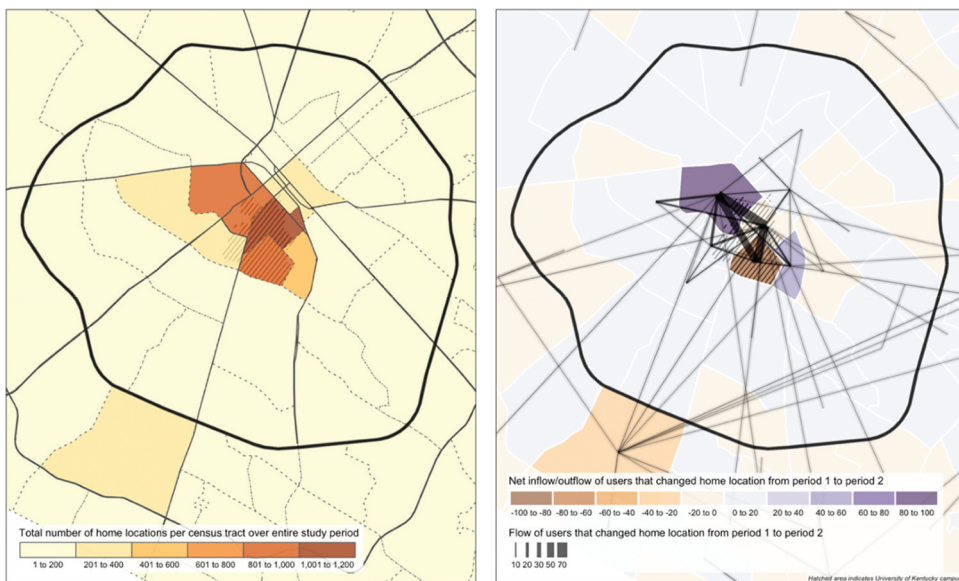


Figure 2. Home locations of Twitter users (a) the total number of home locations per census tract for the entire study period (2012–2017); (b) the number of users whose home locations shifted from period 1 (2012–13) to period 2 (2014–2017). Lines represent these movements, while shading in the choropleth map indicates net inflow/outflow.⁵

Indeed, most of the net changes (either positive or negative) in home location seen in [Figure 2](#) were concentrated in and between these areas around the University of Kentucky campus, pointing to the centrality of these places within the digital expression of social life in Lexington. Other tracts with a relative high number of visits are the downtown core of Lexington (northeast of the UK campus), a suburban neighborhood in the south containing Fayette Mall, the city's largest indoor shopping mall, and a similar neighborhood in the city's east containing Hamburg Place, a series of interconnected strip malls and big box stores. This activity pattern is consistent with what one would expect for an American city with a large university during this time period. However, of greater interest in relation to gentrification is whether and how this pattern changes over time.

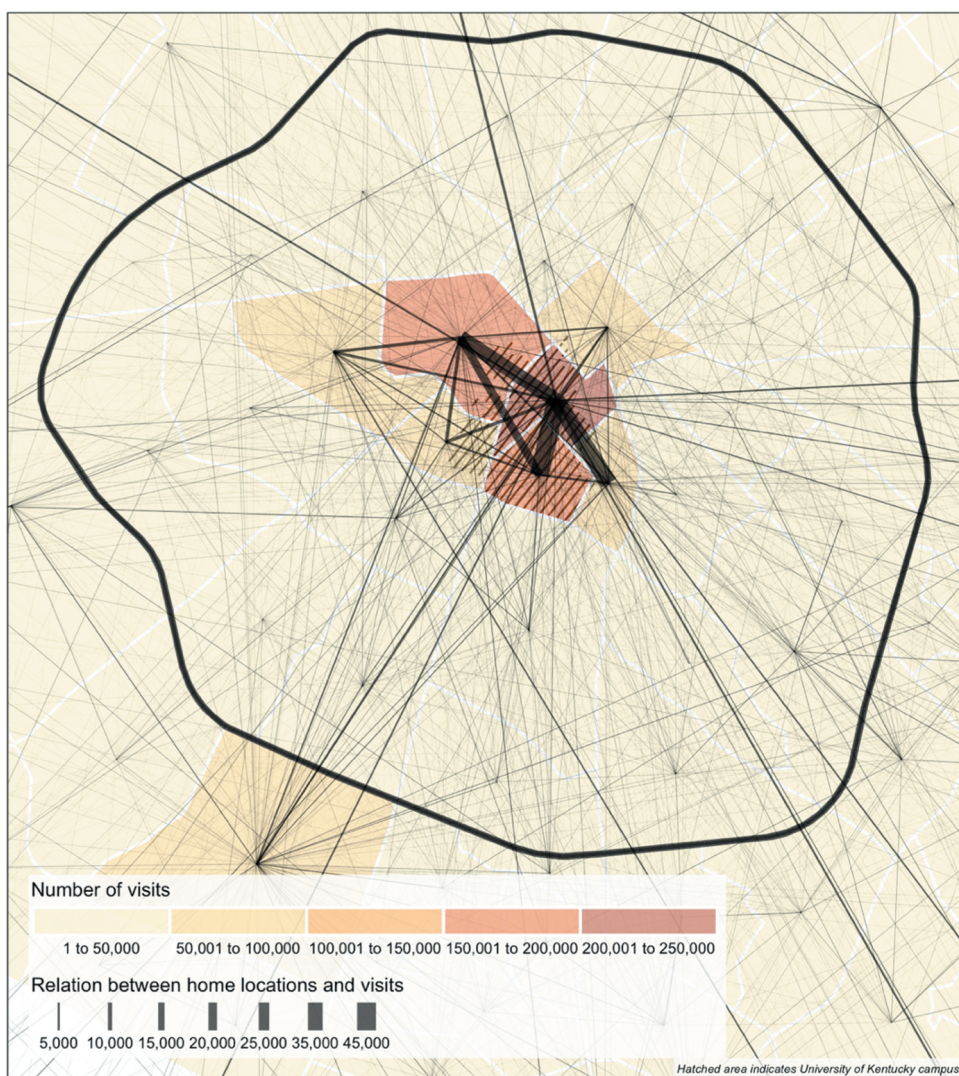


Figure 3. Connections between users' home location and visits to other census tracts. Lines represent visits; the shading in the choropleth layer indicates the total number of visits in each census tract.

To address that question, Figure 4 illustrates the same connections between home locations and visits shown in Figure 3 but for the two time periods.⁶ At the macro-scale, no apparent large shift in the network of connections occur – except for a relative increase in the visits to Fayette Mall in the south and a greater density of connections between the core neighborhoods around the university. On the surface, the patterns of mobility and visits of Twitter users remain relatively stable over the five-plus years of the study period. This echoes findings elsewhere in the literature that neighborhoods mostly stay the same over time rather than experiencing drastic change (Solari, 2012; Cortright & Mahmoudi, 2014; Delmelle, 2017; Malone & Redfearn, 2018; Connor et al., 2020; Kinahan, 2020). In this case, however, our analysis shows that this stability in neighborhood character not only applies to the *internal* characteristics of neighborhoods, but also to how neighborhoods connect to each other relationally.

To gain another perspective on this, we also examine the share of activity by visitors relative to locals (i.e. users whose home location is the same tract in which a tweet is sent). Thus, in Figure 5, the darkest shaded tracts represent places where more than 80% of tweets are sent by visitors. Looking at the relative share of visitors in each tract again reveals a relatively stable pattern over the two time periods (Figures 5a and 5b). The downtown core and shopping venues in the southern and eastern parts of the city have the highest ratio of tweets from visitors versus locals (the darkest shading in map), while largely residential neighborhoods generally see fewer visits from non-locals.

Gentrification on the Northside of Lexington

These macro-scale patterns, however, provide only limited insight on neighborhood-level gentrification processes within Lexington. To address this, we focus on a specific

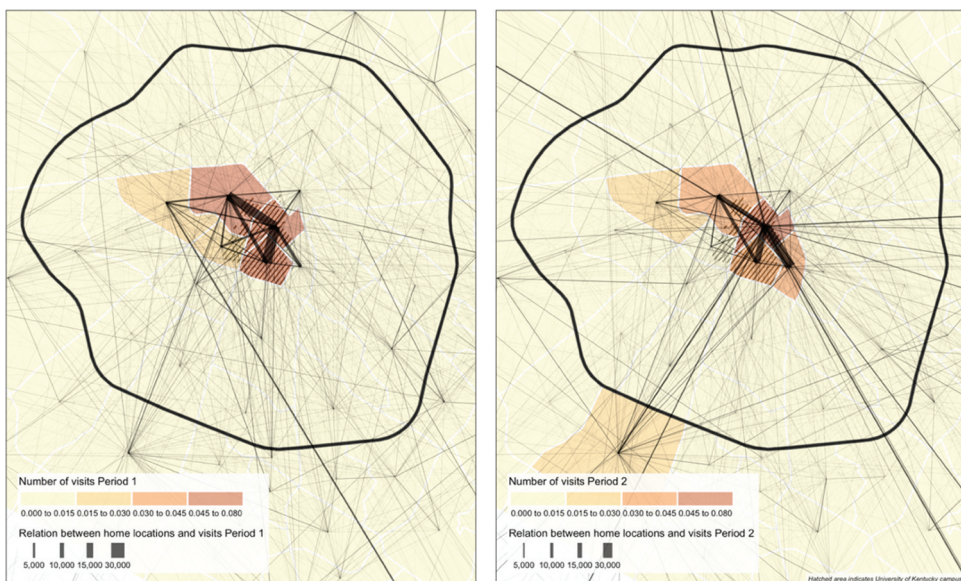


Figure 4. Connections between users' home location and visits to other census tracts (a) June 2012 to December 2013 and (b) January 2014 to December 2017.

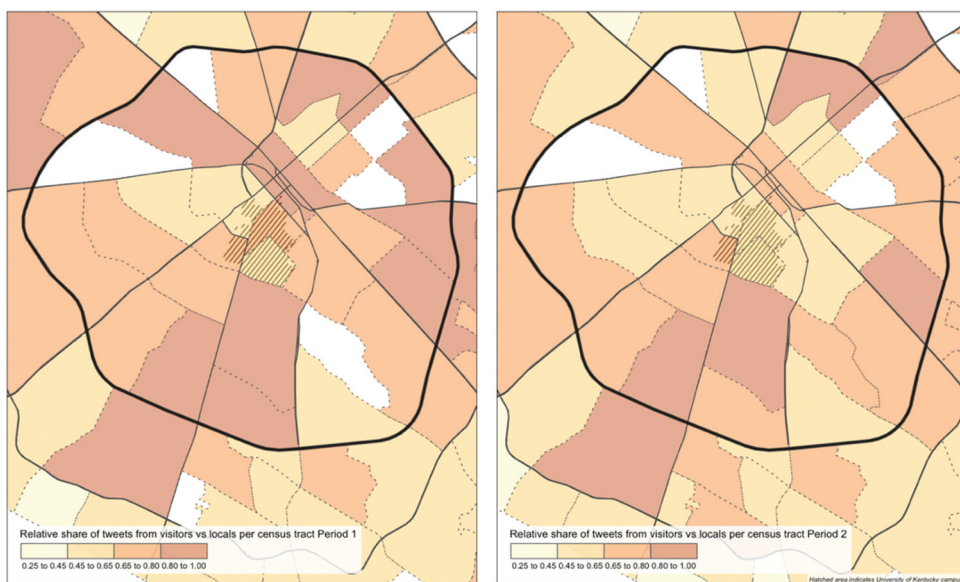


Figure 5. The relative share of tweets from visitors vs. tweets from locals per census tract (a) June 2012 to December 2013 and (b) January 2014 to December 2017.

neighborhood known as the Northside, located just north of downtown (see [Figure 1](#)), which has undergone considerable change over the course of our 2012–2017 study period. As is generally the case with gentrification, delineating specific spatial and temporal boundaries for analysis is not straightforward, as gentrification does not unfold in clear and spatially contiguous units. In the Northside, two parallel commercial corridors – Jefferson Street and North Limestone Avenue– on opposite sides of a major transportation artery, Broadway Avenue, have been the sites of a number of private redevelopment efforts. Beginning with the purchase of two longstanding Northside establishments – Stella’s Kentucky Deli on Jefferson Street in 2006 and Al’s Bar on North Limestone in 2007 – capital investment began to flood into these areas, with new restaurants, bars and other commercial establishments serving as the foundation on which residential real estate speculation took off in surrounding neighborhoods. Even several years ago, a relatively small number of landlords and developers controlled hundreds of residential units in the neighborhood (Lexington Housing Studies, [2015](#)). This resulted in newly-renovated one bedroom apartments being rented for more than double the overall area median rent and houses flipped for 10 times more than their sale price just two years prior (Lexington Housing Studies, [2016](#)).

[Figure 6](#) outlines the locations of the Limestone and Jefferson corridors (demarcated by the hashed area) including key sites of gentrification-related, consumption-oriented attractions (e.g. bars and coffeeshops). Also highlighted is the concentration of Twitter activity relative to the surrounding residential areas and consumption venues. While the point data from Twitter makes it possible to use these areal definitions for the Northside neighborhood, we instead focus on two specific Census tracts – Census tracts 2 and 3 indicated by thick, black lines in [Figure 6](#) – in order to link to demographic data from the US Census Bureau’s American Community Survey. As [Figure 6](#) shows, the vernacular



Figure 6. Map of the Limestone Avenue and Jefferson Street Corridors (delineated with dashed lines). Thicker solid lines indicate Census tracts 2 and 3 used in our analysis to represent the Northside neighborhood. Dots represent a 5% sample of tweets in the dataset.

understanding of neighborhoods and key sites of gentrification does not necessarily align with statistical geographies, as the Jefferson Street Corridor extends significantly into Census tract 1.02 to the south. However, this tract also includes Rupp Arena, home of the Kentucky Wildcats basketball team, which is a popular (tweeting) destination, and part of a very different, distinct downtown neighborhood.

Figure 7a highlights the Northside Census tracts and shows that, relative to other tracts in the urban core of Lexington, the Northside has a modest percentage of visitors during the overall time period of our study (~60%). Based on our contextual understanding of this neighborhood, we would expect to find an increase in the share of visitors over the years of the study period. However, Figure 7b shows that the percentage of visitors is very stable for both tracts. This runs counter to our initial expectations and warrants a closer investigation.

To do this, we shift to examining the *origin* of visitors more closely. Figure 8 demonstrates that the home locations of visitors change quite noticeably between the two periods. For example, in the second period, the number of visitors from the tracts around the University of Kentucky increased substantially, suggesting that the characteristics of those visitors changed in parallel. Drawing upon the existing gentrification literature, we expect the visitor profile for a gentrifying neighborhood to be (come) more highly educated, whiter and richer. In the subsequent section, we consider this more closely by constructing *profiles for each visitor* using education, race, and income characteristics of their home location.⁷

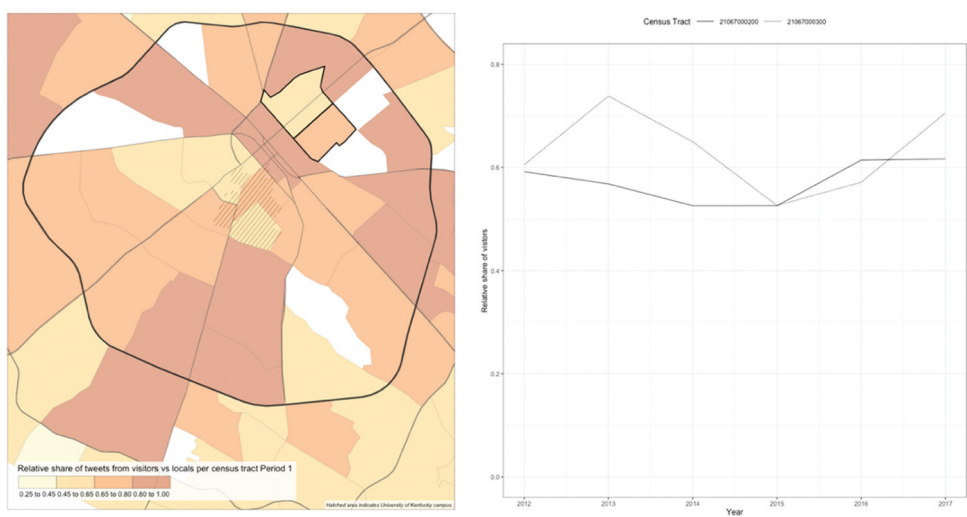


Figure 7. (a) The relative share of tweets from visitors vs locals (Northside census tracts indicated by thicker lines) (b) Percentage of visitors to the two Northside tracts year-over-year

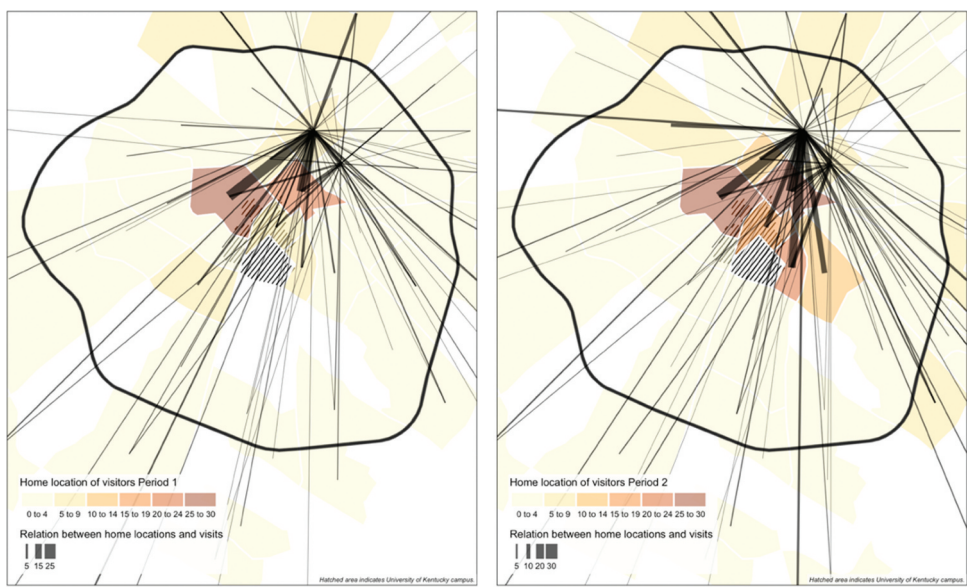
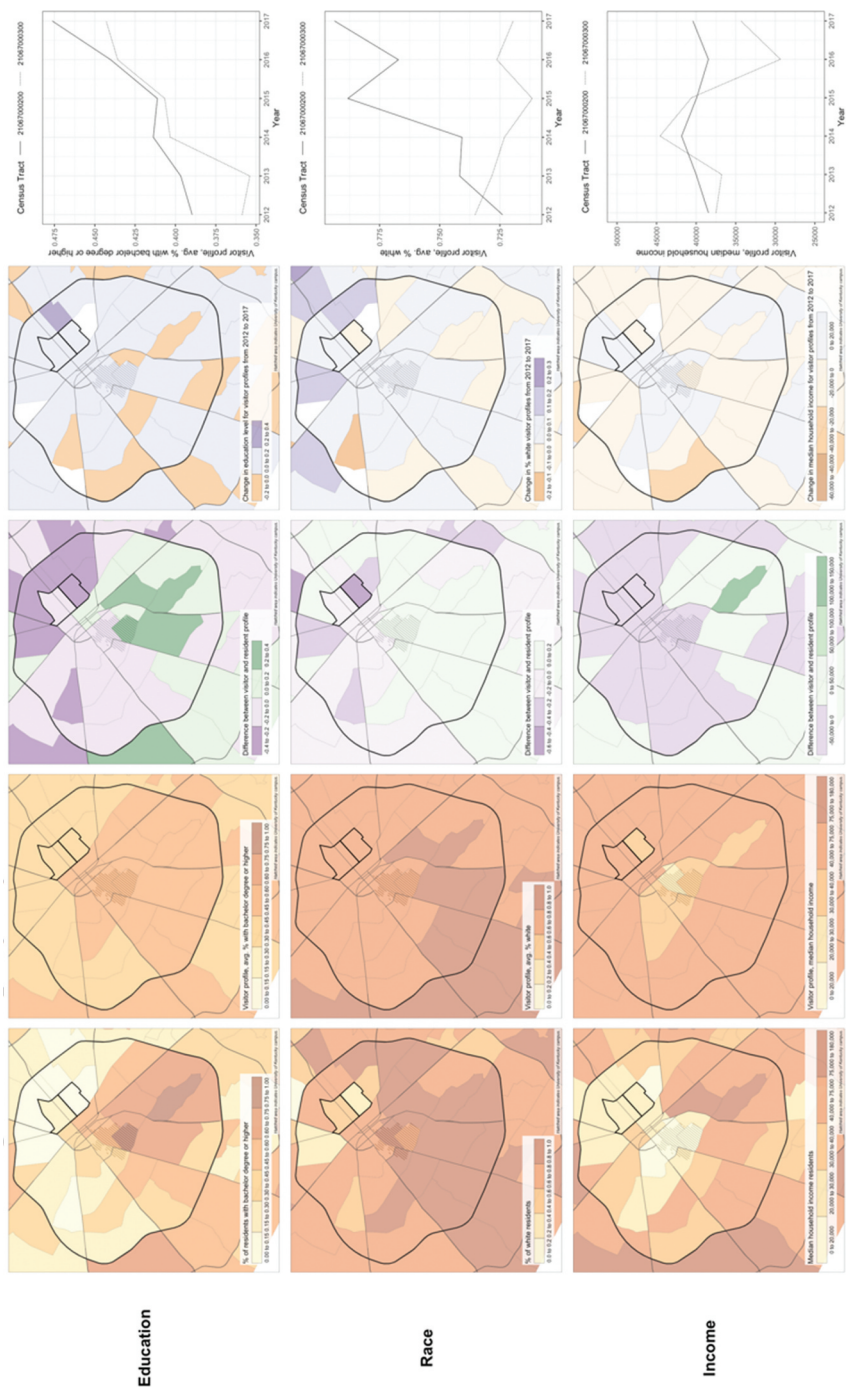


Figure 8. Source of visitors to Northside (a) June 2012 to December 2013 and (b) January 2014 to December 2017.

Changing demographics of visitors to Lexington’s northside

While determining the demographic characteristics of visitors to the Northside is extremely fraught, using the previously identified home locations allows us to make broad inferences about the kinds of people that visit the Northside before and after the neighborhood’s gentrification process intensified. Drawing from the existing literature on using population-level statistics to measure gentrification, we focus on three variables:



educational attainment (as measured by the percentage of residents from a given tract with a bachelor's degree or higher), race (as measured by the percentage of residents who are white) and class (as measured by the median household income).

Figure 9 represents a composite image of how these three variables manifest across the Lexington landscape (leftmost column), the average profile of visitors to each Census tract based on their home Census tract (second column), the difference between these two (third column), the change in visitor profiles between the study's two time periods (fourth column) and a chart depicting the change in average visitor profiles for Census tracts 2 and 3 over the six years of the study period.

Looking at the leftmost column of Figure 9 across all three variables, it is evident that the Northside neighborhood has lower educational levels, lower proportions of white residents and lower median household incomes than other neighborhoods in Lexington's urban core. This reflects the area's legacy as a marginalized, predominantly black neighborhood. In contrast, the second column shows that the average profile of visitors to the Northside neighborhood is considerably better educated, whiter and wealthier than the neighborhood itself. This difference is most visible in the third column of Figure 9, where the purple shading signifies higher values across each of the three variables for the average visitor as compared to the neighborhood.

There is, however, some intra-neighborhood difference in how these inequalities manifest within the Northside. For instance, in Tract 2 (containing the Jefferson Street corridor to the northwest) we find smaller differences between locals and visitors in terms of educational attainment (18% to 26%) and race (16% to 42%) compared to Tract 3 (containing North Limestone), but somewhat larger differences in terms of class (\$17,700 to \$12,700).

The fourth and fifth columns of Figure 9 compare changes in the average visitor profiles for each tract from 2012 through 2017. Our expectation that, as the Northside continues to gentrify, the average profile of visitors will become more educated, whiter and wealthier, is shown to generally be true, with some important caveats. With regards to changes in educational attainment over time, the average visitor profile sees a consistent 8 percentage point increase in the proportion with a bachelor's degree or higher across both Census tracts. But across both the racial and income variables, the two Northside tracts experience somewhat different and inconsistent patterns in changes in the average profile of visitors over time. While there is a general upward trend in the proportion of white visitors for Tract 2 amounting to a 5 percentage point increase from 2012–2017, the racial makeup of visitors to Tract 3 remains fairly stable over time with a slight decrease by the end of the study period.

Similarly, the chart in the bottom-right corner of Figure 9 shows that the average income of visitors to the Northside does not increase substantially but is instead quite stable throughout our study period. Given the prominent role that class and income play in theories of gentrification, these results are somewhat confounding. But rather than taking this as evidence that gentrification *isn't* occurring in the Northside, we instead argue that this highlights precisely the complexity of gentrification as an economic and cultural process, especially in relation to how class is disguised in demographic variables. As noted in our conceptual review earlier, a solitary income indicator misses the complex ways in which class is manifest. For example, young people in the early stages of their careers (who are often identified as harbingers of gentrification) may currently have low

incomes, but over their lifetime will transition into higher income positions (see Freeman, 2005; Hammel & Wyly, 1996; Schuler et al., 1992). Thus, the stability of income levels in Northside visitors, combined with the changes in the educational profiles during our study period are therefore consistent with the findings of earlier gentrification research regarding the primacy of educational attainment as a statistical indicator of changing class compositions in urban neighborhoods.

IV. Conclusion

This analysis provides a proof-of-concept of how big data, specifically geotagged social media, may be leveraged to gain a relational perspective on the process of gentrification. While research on neighborhood change in general, and gentrification in particular, has a long-standing interest in identifying specific pathways or trajectories of change (Delmelle, 2017), studies on such trajectories have been limited mostly to studying the changing characteristics of *residents*. In contrast, our approach examines neighborhood change by studying the evolving characteristics of *visitors* and the changing social relations these visitors engender.

Shifting from data based on residents, an artifact of official data regimes, our approach focuses on the mobility of individuals and the relation between residential location and their activity space. Through this, we are able to document a different view of neighborhood change that reflects changes in use and presence in an area. In the case of Lexington's Northside – a neighborhood undergoing substantial urban change during our study period – this manifests through (an increasing) difference in the characteristics and composition of visitors relative to residents. In short, visitors to the Northside over time increasingly came from whiter and more highly-educated neighborhoods.

But in addition to demonstrating how we might operationalize relational theories of urban space within gentrification studies, the adoption of big data highlights other potential avenues for further exploration. For example, this analysis still relies upon Census tract geographies, but there is nothing inherently preventing a more flexible definition of neighborhoods given that this type of data is often point based. While it remains useful to associate data points to census tracts to enrich it with Census and other preexisting datasets, as we do in this paper, alternative approaches are also possible. For example, we could have used social media data to determine users' home census tracts in order to construct the *profile* of visitors, but use our own neighborhood boundaries for *visits*. This could help further address the spatial granularity and "fit" issue highlighted in the discussion of the Jefferson and Limestone corridors (cf. Figure 6). Additionally, although we have focused on Lexington in this study, the widespread availability of social media and other types of big data opens the door for more comparative analyses of gentrification across a larger number of cities.

That said, big data approaches are hardly a magic bullet for all the issues facing urban research and moreover, this approach brings its own new challenges. First, there is increasing concern around the significant ethical and privacy implications of the existence of this data in public discourse and its use in research (e.g. boyd and Crawford, 2012; Elwood & Leszczynski, 2011; Taylor, 2016; Zook et al., 2017). Many of these questions and implications are especially relevant here since analyzing neighborhood change requires data at a very fine temporal and spatial scale, which can be directly at

odds with the privacy of and potential harm to the people included in these datasets. While the Census has clear-cut rules and is highly restrictive toward access to micro-data sets, many big data sources, often from private companies, do not have similar safeguards.

Furthermore, rising awareness of privacy implications has resulted in severe restrictions on the public availability of data for research in recent years. While this can be seen as a positive change, it also risks creating situations where only researchers embedded within big data companies can access data. Given the profit-maximizing incentives of private companies, this will likely yield different questions and applications than those asked by academic researchers, e.g. the construction of analytical toolkits for gentrification hotspot prediction to aid real estate speculation. As such, regulations for the ethical use and sharing of such data are needed to make big data a sustainable addition to gentrification and other urban social research.

Many big data sets are by-products of platforms, products and processes that are designed with a different purpose than what such data might ultimately be used for. Most certainly, such data differs from conventional survey or census data sets designed to be representative of a larger population. While the inherent bias that this introduces does not invalidate its potential use for social science research, a continuous and rigorous evaluation of representation and bias is warranted if we want to use such data for social science and policy input beyond the current proof-of-concept stage.

Moreover, it is important that researchers do not limit themselves to a single big data source as a cure-all. Social media platforms and related technologies tend to have limited lifespans and/or may only be used by specific population groups, meaning that such data will never be able to replace Census, survey and register-based datasets. In other words, social media data are not simply a cheaper and better alternative to official datasets. Rather their use should be targeted to fill specific gaps in existing approaches such as our focus on building a relational understanding of gentrification. Even in this analysis, we rely heavily on both the American Community Survey and our own qualitative and experiential knowledge to contextualize and enrich what can be inferred from social media data alone.

Ultimately, we argue that big data is not a panacea for empirical studies of gentrification, or for any particular urban issue of interest, and given the “multidimensionality of gentrification” still means that “the use of a single variable to identify it is almost certain to fail” (Bostic & Martin, 2003, p. 2431). That said, our analysis has demonstrated that despite being *unable* to capture some key aspects of the gentrification process, the spatial and temporal granularity of social media data *can* supplement our understandings of where gentrification is occurring, and how gentrification is a fundamentally relational process that links different spaces together through people’s everyday mobility and consumption patterns. Indeed, the real-time nature of this data offers an opportunity to help serve as one aspect of an “early warning system” for gentrification (cf. Chapple & Zuk, 2016), identifying how changing urban mobility patterns signal broader changes within the urban fabric.

Notes

1. Specifically, a bounding box with 37.67372° N 85.08614° W as the southwest corner and 38.38294° N 83.92788° W as the northeast corner. This includes the nearby towns of Frankfort to the west, Georgetown and Paris to the north and Harrodsburg and Richmond in the south.
2. To improve readability of the paper we do not use quotes around home for the rest of the article although our intent remains the same.
3. Geotagged tweets generally have a relatively precise point location (stored as longitude/latitude) attached to them which allows them to be easily aggregated to larger and flexibly defined spatial units that conform to the lived experience of neighborhoods. While this been used successfully at a range of scales in previous work (cf. Poorthuis, 2018; Poorthuis et al., 2020; Shelton et al., 2015) the analysis in this paper uses conventional Census tracts as aggregation units in order to more easily connect with data from the American Community Survey.
4. These periods do not have the same length as they have been chosen to divide the dataset equally in terms of total number of tweets. Because geotagged tweeting became less prevalent from mid-2015 onwards, the latter period is longer. Defining time periods in this way also allows us to distinguish between the period prior to the end of 2013 when many key commercial establishments first opened (West Sixth Brewery, Arcadium, North Lime Coffee and Donuts) and from 2014 onward as they received more visitors.
5. Data spans the entire metropolitan region of Lexington, KY but figures zoom in on Lexington's urban core for clarity and simplicity.
6. Because the two time periods are of different length, we do not use absolute counts. Rather we use the relative frequency of visits (Number of Visits to Tract_{*i*}/Total Number of Visits to All Tracts) to make the periods comparable.
7. These characteristics are derived from the American Community Survey 2012–2016 5-year estimates (as to overlap with our study period).

Acknowledgments

The authors would like to thank Qingqing Chen for her assistance with preliminary data analysis for this article.

Data Availability Statement

The data and code that support the findings of this study are available on <https://github.com/atepoorthuis/relationalgeographiesofgentrification> and <http://doi.org/10.5281/zenodo.4535943>.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ate Poorthuis  <http://orcid.org/0000-0002-3808-7493>
 Taylor Shelton  <http://orcid.org/0000-0003-0776-9296>
 Matthew Zook  <http://orcid.org/0000-0002-6034-3262>

References

- Atkinson, Rowland. 2000. "Measuring gentrification and displacement in greater london." *Urban Studies* 37 (1): 149–165. <https://doi.org/10.1080/0042098002339>
- Barnes, Trevor J. 2013. "Big data, little history." *Dialogues in Human Geography* 3 (3): 297–302. <https://doi.org/10.1177/2043820613514323>
- Barnes, Trevor J., & Wilson, Matthew W. 2014. "Big data, social physics, and spatial analysis: The early years." *Big Data & Society* 1 (1): 1. <https://doi.org/10.1177/2053951714535365>
- Barton, Michael. 2016. "An exploration of the importance of the strategy used to identify gentrification." *Urban Studies* 53 (1): 92–111. <https://doi.org/10.1177/0042098014561723>
- Beekmans, Jeroen. 2011. "Check-In Urbanism: Exploring Gentrification through Foursquare Activity." MSc Thesis, University of Amsterdam.
- Blanford, Justine I., Huang, Zhuojie, Savelyev, Alexander, & MacEachren, Alan M. 2015. "Geo-located tweets. enhancing mobility maps and capturing cross-border movement." *Plos One* 10 (6): e0129202. <https://doi.org/10.1371/journal.pone.0129202>
- Bostic, Raphael W., & Martin, Richard W. 2003. "Black home-owners as a gentrifying force? Neighbourhood dynamics in the context of minority home-ownership." *Urban Studies* 40 (12): 2427–2449. <https://doi.org/10.1080/0042098032000136147>
- boyd, danah, and Kate Crawford. 2012. "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." *Information, Communication & Society* 15 (5): 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Buntin, John. 2015. "The myth of gentrification." *Slate*. 14 January. Available from: http://www.slate.com/articles/news_and_politics/politics/2015/01/the_gentrification_myth_it_s_rare_and_not_as_bad_for_the_poor_as_people.html
- Carlson, H. J. 2020. "Measuring displacement: Assessing proxies for involuntary residential mobility." *City & Community*. 19 3 573–592 <https://doi.org/10.1111/cico.12482>
- Chapple, Karen, & Zuk, Miriam. 2016. "Forewarned: The use of neighborhood early warning systems for gentrification and displacement." *Cityscape* 18(3): 109–130. <https://www.jstor.org/stable/26328275>
- Chen, Qingqing, & Poorthuis, Ate. 2021. "Identifying home locations in human mobility data: An open-source R package for comparison and reproducibility." *International Journal of Geographical Information Science*. In press. doi: 10.1080/13658816.2021.1887489
- Connor, Dylan Shane, Gutmann, Myron P., Cunningham, Angela R., Clement, Kerri Keller, & Leyk, Stefan. 2020. "How entrenched is the spatial structure of inequality in cities? Evidence from the integration of census and housing data for denver from 1940 to 2016." *Annals of the American Association of Geographers* 110 (4): 1022–1039. <https://doi.org/10.1080/24694452.2019.1667218>
- Cortright, Joe. 2015. "In Defense of gentrification." *The Atlantic*. 31 October. Available from: <http://www.theatlantic.com/business/archive/2015/10/in-defense-of-gentrification/413425/>
- Cortright, Joe, & Mahmoudi, Dillon. 2014. "Lost in Place: Why the persistence and spread of concentrated poverty—not gentrification—is our biggest urban challenge." *City Observatory*. Available from: <http://cityobservatory.org/lost-in-place/>
- Delmelle, Elizabeth C. 2017. "Differentiating pathways of neighborhood change in 50 US metropolitan areas." *Environment & Planning A* 49 (10): 2402–2424. <https://doi.org/10.1177/0308518X17722564>
- Easton, Sue, Lees, Loretta, Hubbard, Phil, & Tate, Nicholas. 2020. "Measuring and mapping displacement: The problem of quantification in the battle against gentrification." *Urban Studies* 57 (2): 286–306. <https://doi.org/10.1177/0042098019851953>
- Economist, The. 2015. "Bring on the hipsters: Gentrification is good for the poor." *The Economist*. 19 February. Available from: <https://www.economist.com/news/united-states/21644164-gentrification-good-poor-bring-hipsters>
- Elwood, Sarah, & Leszczynski, Agnieszka. 2011. "Privacy, reconsidered: New representations, data practices, and the geoweb." *Geoforum* 42 (1): 6–15. <https://doi.org/10.1016/j.geoforum.2010.08.003>

- Fairfield, John D. 1994. "The scientific management of urban space: Professional city planning and the legacy of progressive reform." *Journal of Urban History* 20 (2): 179–204. <https://doi.org/10.1177/009614429402000202>
- Florida, Richard. 2014. "No one's very good at identifying gentrification." *CityLab*. 15 December. Available from: <https://www.citylab.com/equity/2014/12/no-ones-very-good-at-correctly-identifying-gentrification/383724/>
- Ford, George. 1913. "The city scientific." *Engineering Record* 67: 551–552.
- Freeman, Lance. 2005. "Displacement or succession? Residential mobility in gentrifying neighborhoods." *Urban Affairs Review* 40 (4): 463–491. <https://doi.org/10.1177/1078087404273341>
- Glass, Ruth. 1964. *London: Aspects of Change*. MacGibbon & Kee.
- Hammel, Daniel J., & Wyly, Elvin K. 1996. "A model for identifying gentrified areas with census data." *Urban Geography* 17 (3): 248–268. <https://doi.org/10.2747/0272-3638.17.3.248>
- Hargittai, Eszter. 2015. "Is bigger always better? Potential biases of big data derived from social network sites." *The Annals of the American Academy of Political and Social Science* 659 (1): 63–76. <https://doi.org/10.1177/0002716215570866>
- Hwang, Jackelyn, & Sampson, Robert J. 2014. "Divergent pathways of gentrification racial inequality and the social order of renewal in Chicago neighborhoods." *American Sociological Review* 79 (4): 726–751. <https://doi.org/10.1177/0003122414535774>
- Kinahan, Kelly L. 2020. "Spatial–Temporal Neighborhood Patterns in Four Legacy Cities." *Urban Affairs Review* 57 (2): 492–526. <https://doi.org/10.1177/1078087419897825>
- Kirkland, Elizabeth. 2008. "What's race got to do with it? Looking for the racial dimensions of gentrification." *Western Journal of Black Studies* 32 (2): 18–30.
- Kreager, Derek A., Lyons, Christopher J., & Hays, Zachary R. 2011. "Urban revitalization and Seattle crime, 1982–2000." *Social Problems* 58 (4): 615–639. <https://doi.org/10.1525/sp.2011.58.4.615>
- Kwan, Mei-Po. 2012. "The uncertain geographic context problem." *Annals of the Association of American Geographers* 102 (5): 958–968. <https://doi.org/10.1080/00045608.2012.687349>
- Kwan, Mei-Po. 2013. "Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility." *Annals of the Association of American Geographers* 103 (5): 1078–1086. <https://doi.org/10.1080/00045608.2013.792177>
- Lawton, Philip. 2020. "Unbounding gentrification theory: Multidimensional space, networks and relational approaches." *Regional Studies* 54 (2): 268–279. <https://doi.org/10.1080/00343404.2019.1646902>
- Lefebvre, Henri. 1991. *The Production of Space*. Blackwell: Oxford.
- LeGates, Richard, Tate, Nicholas J., & Kingston, Richard. 2009. "Spatial thinking and scientific urban planning." *Environment and Planning, B, Planning & Design* 36 (5): 763–768. <https://doi.org/10.1068/b3605com>
- Lexington Housing Studies. 2015. "Who's gentrifying northeast Lexington?" Lexington housing studies blog. 15 October. Available from: <https://medium.com/@HousingLEX/who-s-gentrifying-northeast-lexington-7144f80066e8>
- Lexington Housing Studies. 2016. "Dispelling the myths of Lexington's gentrification apologists." *Lexington Housing Studies Blog*. 19 October. Available from: <https://medium.com/@HousingLEX/dispelling-the-myths-of-lexingtons-gentrification-apologists-f259ab2eea02>
- Light, Jennifer S. 2003. *From warfare to welfare: Defense intellectuals and urban problems in cold war America*. Johns Hopkins University Press.
- Luo, Feixiong, Cao, Guofeng, Mulligan, Kevin, & Xiang, Li. 2016. "Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago." *Applied Geography* 70: 11–25. <https://doi.org/10.1016/j.apgeog.2016.03.001>
- Madden, David J. 2018. "Pushed off the map: Toponymy and the politics of place in New York City." *Urban Studies* 55 (8): 1599–1614. <https://doi.org/10.1177/0042098017700588>
- Malone, Thom, & Redfearn, Christian L. 2018. "Shocks & ossification: The durable hierarchy of neighborhoods in U.S. metropolitan areas from 1970 to 2010." *Regional Science and Urban Economics* 69: 94–121. <https://doi.org/10.1016/j.regsciurbeco.2018.01.002>

- Papachristos, Andrew V., Smith, Chris M., Scherer, Mary L., & Fugiero, Melissa A. 2011. "More coffee, less crime? The relationship between gentrification and neighborhood crime rates in Chicago, 1991 to 2005." *City & Community* 10 (3): 215–240. <https://doi.org/10.1111/j.1540-6040.2011.01371.x>
- Patrick, Heidkamp, Patrick, C., & Lucas, Susan. 2006. "Finding the gentrification frontier using census data: The case of Portland, Maine." *Urban Geography* 27 (2): 101–125. <https://doi.org/10.2747/0272-3638.27.2.101>
- Phillips, Nolan E., Levy, Brian L., Sampson, Robert J., Small, Mario L., & Wang, Ryan Q. 2019. "The social integration of American cities: Network measures of connectedness based on everyday mobility across neighborhoods." *Sociological Methods & Research*. 004912411985238 <https://doi.org/10.1177/0049124119852386>
- Poorthuis, Ate. 2018. "How to draw a neighborhood? The potential of big data, regionalization, and community detection for understanding the heterogeneous nature of urban neighborhoods." *Geographical Analysis* 50 (2): 182–203. <https://doi.org/10.1111/gean.12143>
- Poorthuis, Ate, Power, Dominic, & Zook, Matthew. 2020. "Attentional social media: Mapping the spaces and networks of the fashion industry." *Annals of the American Association of Geographers* 110 (4): 941–966. <https://doi.org/10.1080/24694452.2019.1664887>
- Poorthuis, Ate, & Zook, Matthew. 2017. "making big data small: Strategies to expand urban and geographical research using social media." *Journal of Urban Technology* 24 (4): 115–135. <https://doi.org/10.1080/10630732.2017.1335153>
- Poorthuis, Ate, & Zook, Matthew. 2020. "Being smarter about space: Drawing lessons from spatial science." *Annals of the American Association of Geographers* 110 (2): 349–359. <https://doi.org/10.1080/24694452.2019.1674630>
- Rose, Damaris. 1984. "Rethinking gentrification: Beyond the uneven development of Marxist urban theory." *Environment and Planning, D, Society & Space* 2 (1): 47–74. <https://doi.org/10.1068/d020047>
- Schaefer, Bryan. 2014. "Social Media to Locate Urban Displacement: Assessing the Risk of Displacement Using Volunteered Geographic Information in the City of Los Angeles." M. S. thesis, University of Southern California.
- Schuler, Jill R., Kent, Robert B., & Monroe, Charles B. 1992. "Neighborhood gentrification: A discriminant analysis of a historic district in Cleveland, Ohio." *Urban Geography* 13 (1): 49–67. <https://doi.org/10.2747/0272-3638.13.1.49>
- Shelton, Taylor. 2017. "The urban geographical imagination in the age of big data." *Big Data & Society* 4 (1). 1 <https://doi.org/10.1177/2053951716665129>
- Shelton, Taylor, & Poorthuis, Ate. 2019. "The nature of neighborhoods: Using big data to rethink the geographies of Atlanta's neighborhood planning unit system." *Annals of the American Association of Geographers* 109(5): 1341–1361. <https://doi.org/10.1080/24694452.2019.1571895>
- Shelton, Taylor, Poorthuis, Ate, & Zook, Matthew. 2015. "Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information." *Landscape and Urban Planning* 142: 198–211. <https://doi.org/10.1016/j.landurbplan.2015.02.020>
- Slater, Tom, Curran, Winifred, & Lees, Loretta. 2004. "Gentrification research: New directions and critical scholarship." *Environment & Planning A* 36 (7): 1141–1150. <https://doi.org/10.1068/a3718>
- Smith, Chris M. 2014. "The influence of gentrification on gang homicides in Chicago neighborhoods, 1994 to 2005." *Crime and Delinquency* 60 (4): 569–591. <https://doi.org/10.1177/001128712446052>
- Smith, Neil. 1979. "Toward a theory of gentrification a back to the city movement by capital, not people." *Journal of the American planning association* 45 (4): 538–548.
- Soja, Edward W. 1980. "The socio-spatial dialectic." *Annals of the Association of American Geographers* 70 (2): 207–225
- Solari, Claudia D. 2012. "Affluent neighborhood persistence and change in U.S. Cities." *City & Community* 11 (4): 370–388. <https://doi.org/10.1111/j.1540-6040.2012.01422.x>

- Taylor, Linnet. 2016. "No place to hide? The ethics and analytics of tracking mobility using mobile phone data." *Environment and Planning. D, Society & Space* 34 (2): 319–336. <https://doi.org/10.1177/0263775815608851>
- Twilley, Nicola. 2009. "Cupcake Gentrification". *Edible Geography blog*. 5 August. Available from: <http://www.ediblegeography.com/cupcakegentrification/>
- Twilley, Nicola. 2011. "Mapping Gangs and Cupcakes." *Edible Geography blog*. 9 March. Available from: <http://www.ediblegeography.com/mapping-gangs-and-cupcakes/>
- Venerandi, Alessandro, Quattrone, Giovanni, Capra, Licia, Quercia, Daniele, & Saez-Trumper, Diego. 2015. "Measuring urban deprivation from user generated content." In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 254–264. Vancouver, BC, Canada.
- Wu, Wenjie, Wang, Jianghao, & Dai, Tianshi. 2016. "the geography of cultural ties and human mobility: Big data in urban contexts." *Annals of the American Association of Geographers* 106 (3): 612–630. <https://doi.org/10.1080/00045608.2015.1121804>
- Wyly, Elvin K., & Hammel, Daniel J. 1998. "Modeling the context and contingency of gentrification." *Journal of Urban Affairs* 20 (3): 303–326. <https://doi.org/10.1111/j.1467-9906.1998.tb00424.x>
- Wyly, Elvin K., & Hammel, Daniel J. 1999. "Islands of decay in seas of renewal: Housing policy and the resurgence of gentrification." *Housing Policy Debate* 10 (4): 711–771. <https://doi.org/10.1080/10511482.1999.9521348>
- Zhang, Weiyang, Derudder, Ben, Wang, Jianghao, Shen, Wei, & Witlox, Frank. 2016. "Using location-based social media to chart the patterns of people moving between cities: The case of weibo-users in the yangtze river delta." *Journal of Urban Technology* 23 (3): 91–111. <https://doi.org/10.1080/10630732.2016.1177259>
- Zook, Matthew. 2017. Crowd-sourcing the smart city: Using big geosocial media metrics in urban governance." *Big Data & Society* 4 (1). 1 <https://doi.org/10.1177/2053951717694384>
- Zook, Matthew, Barocas, Solon, danah, boyd, Kate, Crawford, Emily, Keller, Gangadharan, Seeta Peña, Alyssa, Goodman, Rachelle, Hollander, Barbara A., Koenig, Jacob, Metcalf, Arvind, Narayanan, Alondra, Nelson, & Frank, Pasquale. 2017. "Ten simple rules for responsible big data research." *PLoS Computational Biology* 13 (3). e1005399 <https://doi.org/10.1371/journal.pcbi.1005399>
- Zukin, Sharon, Lindeman, Scarlett, & Hurson, Laurie. 2017. "The omnivore's neighborhood? Online restaurant reviews, race, and gentrification." *Journal of Consumer Culture* 17(3): 459–479. <https://doi.org/10.1177/1469540515611203>